

BAB I

PENDAHULUAN

1.1 Latar Belakang

Teknologi dan telekomunikasi yang semakin berkembang menjadi salah satu kebutuhan tersier yang tanpa disadari telah menjadi kebutuhan primer bagi orang-orang tertentu. Perubahan ini jelas mempengaruhi perkembangan teknologi komunikasi contohnya adalah jejaring sosial. Salah satu jejaring sosial yang hingga saat ini masih aktif dengan jumlah pengguna yang meningkat setiap tahunnya adalah *Twitter*.

Jejaring sosial yang membatasi penggunanya untuk mengirim 140 karakter dalam setiap postingan ini merupakan salah satu jejaring sosial *microblogging* dengan pertumbuhan tercepat sejak 2006. Salah satu keunggulan penggunaan *Twitter* adalah mampu memberikan informasi secara *real-time*. Seringkali, informasi tersebut bersifat ambigu hingga mengakibatkan penerima informasi salah menafsirkan informasi yang diberikan. Oleh karenanya diperlukan sebuah sistem untuk memeriksa kebenaran struktur kalimat berdasarkan *grammar* dan *lexicon* sehingga informasi yang diterima memiliki makna.

Bahasa natural adalah bahasa yang diucapkan, ditulis atau diisyaratkan sebagai bahasa komunikasi untuk komputer dan manusia. Komunikasi ini dalam ilmu komputer dikenal dengan *Natural Processing Language* (NLP). Salah satu pengolahan bahasa natural dalam NLP adalah *parsing*. Terdapat 2 (dua) kategori pada metode *parsing* yaitu *top-down* dan *bottom-up*. *Top-down parsing* bekerja dengan menguraikan sebuah kalimat dengan *constituent* terbesar sampai menghasilkan *constituent* terkecil dalam kalimat yaitu kata. *Bottom-up parsing* bekerja sebaliknya yaitu dengan mengambil satu demi satu kata lalu merangkainya menjadi *constituent* terbesar yaitu kalimat.

Left-corner parsing adalah sebuah metode gabungan dimana prosesnya dimulai secara *bottom-up* dan diakhiri dengan *top-down*. Karena parse yang ada penentuan jenis katanya kebanyakan dilakukan secara

manual, agar mempermudah proses *parsing* peneliti memanfaatkan *POS Tagger*.

POS (Part-Of-Speech) Tag adalah sebuah sistem pelabelan kelas kata yang ada dalam suatu kalimat. Bentuk dari proses ini adalah mengidentifikasi kata sebagai nomina, verba, adjektiva, adverbial dan lain – lain.

Penelitian dalam bidang ini sudah banyak dilakukan, terutama di luar negeri. Sementara di Indonesia sendiri masih jarang dilakukan. Oleh karena itu penulis mencoba melakukan penelitian tentang **“*Parsing Twitter Menggunakan Metode Left-Corner Parsing Dengan Memanfaatkan POS Tagger.*”**

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, maka rumusan masalah yang diangkat dalam penelitian ini adalah :

1. Bagaimana membuat data *training* dari *tweet – tweet* yang menggunakan bahasa Indonesia formal?
2. Bagaimana mengenali kelas kata secara otomatis?
3. Bagaimana menentukan aturan *left-corner parsing*?

1.3 Tujuan Penelitian

Tujuan dari penelitian berdasarkan rumusan masalah di atas adalah :

1. Membuat data *training* dari *tweet-tweet* yang menggunakan bahasa Indonesia formal.
2. Mengenali kelas kata secara otomatis.
3. Menentukan aturan *left-corner parsing*.

1.4 Batasan Masalah

Dalam penelitian ini penulis membatasi masalah dalam konsep dan implementasinya. Di antaranya adalah :

1. Menggunakan *tweet* berbahasa Indonesia dan berbahasa formal.
2. Pengenalan kelas kata menggunakan Stanford *POS Tagger*.

3. Algoritma *parsing* yang digunakan adalah *left-corner parsing*.

1.5 Metodologi

Tahapan – tahapan dalam penulisan tugas akhir ini adalah sebagai berikut :

1.5.1 Studi Pustaka

Tahap ini adalah tentang bagaimana melakukan pemahaman tentang konsep sistem yang akan digunakan serta mempelajari literatur – literatur yang berkaitan. Antara lain adalah *Twitter*, aturan bahasa Indonesia, *POS Tagger* Stanford NLP serta aturan *left-corner parsing*.

1.5.2 Analisa dan Rancangan Sistem

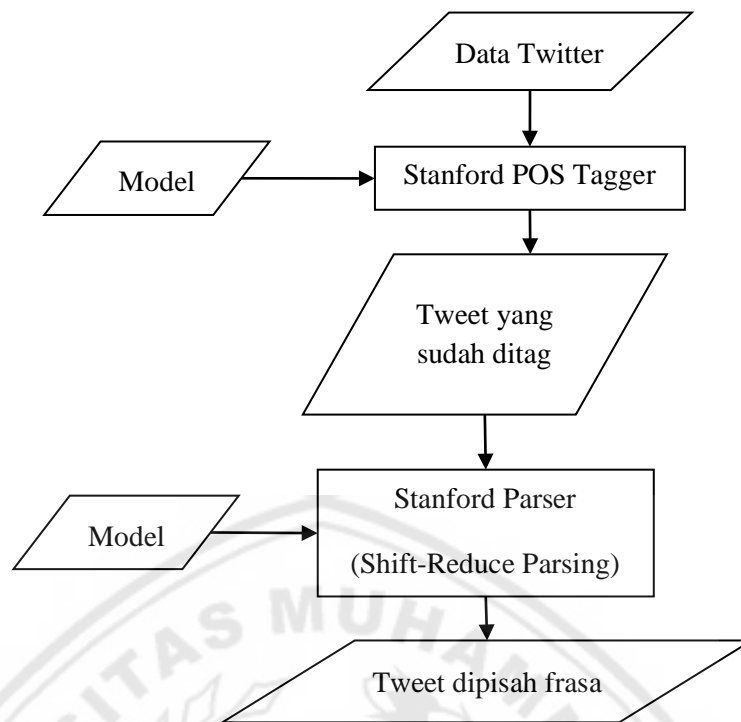
Tahap ini peneliti melakukan analisa dan mendefinisikan kebutuhan apa saja yang dibutuhkan serta merancang sistem untuk tugas akhir ini.

1.5.3 Implementasi Sistem

Tahap ini adalah implementasi sistem dimana dimana peneliti mengimplementasikan sistem yang telah dirancang pada tahap sebelumnya.

1.5.4 Arsitektur Sistem

Tahap ini adalah menggambarkan bagaimana arsitektur sebuah sistem yang akan di bangun. Berikut adalah arsitektur sistem untuk tugas akhir ini:



Gambar 1.1 Arsitektur Sistem

1.5.5 Pengujian

Tahap ini adalah pengujian terhadap sistem yang telah dirancang. Pengujian yang dimaksud adalah pengujian akurasi. Dalam pengujian akurasi ini terdapat 2 (dua) model atau tahap pengujian. Yang pertama adalah pengujian akurasi kata, dan yang berikutnya adalah pengujian akurasi kalimat.

1.5.6 Pembuatan Laporan

Tahap terakhir adalah pembuatan laporan. Dimana laporan yang dibuat harus disesuaikan dengan format yang telah ditentukan. Dimulai dari pembangunan sistem hingga pengujian sistem .

1.6 Sistematika Penulisan

Tugas akhir ini disusun dalam bentuk laporan yang telah dibagi secara sistematika menjadi 5 (lima) bab, sebagai berikut :

BAB I : PENDAHULUAN

Bab I membahas tentang bagaimana latar belakang masalah dari tugas akhir, apa saja rumusan masalahnya, apa tujuan dari penelitian, apa saja batasan masalahnya, metodologi serta sistematika penulisan dalam tugas akhir ini.

BAB II : LANDASAN TEORI

Bab II membahas tentang teori – teori yang berhubungan dengan penyelesaian tugas akhir ini.

BAB III : ANALISA DAN PERANCANGAN SISTEM

Bab III membahas tentang analisa dan perancangan dari sebuah sistem yang di bangun untuk menyelesaikan tugas akhir ini.

BAB IV : IMPLEMENTASI DAN PENGUJIAN

Bab IV membahas tentang hasil dari perancangan serta pengujian yang dilakukan.

BAB V : PENUTUP

Bab V membahas tentang hasil dari keempat tahap sebelumnya. Berisi tentang kesimpulan dan saran untuk pengembangan sistem selanjutnya.